Naval Research Laboratory

Washington, DC 20375-5320

NRL/MR/5650--08-9146

The Feasibility of Nonlinear Dimensionality Reduction for the Rapid Analysis of Persistent Surveillance Data, Including the Detection of IED Placement Activity

Frank Bucholtz

Photonics Technology Branch
Optical Sciences Division

Jonathan M. Nichols

Optical Techniques Branch
Optical Sciences Division

Michael D. Duncan
Leslie N. Smith

Applied Optics Branch
Optical Sciences Division

October 31, 2008

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
31-10-2008

2. REPORT TYPE
Memorandum Report

3. DATES COVERED *(From - To)*
15 April 2007 – 30 Sep 2008

4. TITLE AND SUBTITLE
The Feasibility of Nonlinear Dimensionality Reduction for the Rapid Analysis of Persistent Surveillance Data, Including the Detection of IED Placement Activity

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
ONR 61135N

6. AUTHOR(S)
Frank Bucholtz, Jonathan M. Nichols, Michael D. Duncan, and Leslie N. Smith

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
WU EW014-02-41-8955

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/MR/5650--08-9146

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
One Liberty Center
875 North Randolph Street
Arlington, VA 22203-1995

10. SPONSOR / MONITOR'S ACRONYM(S)
ONR

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Our ability to analyze large, complex data sets, such as persistent surveillance data, has often far outstripped our ability to rapidly analyze that data. We have identified a class of intelligent data reduction algorithms, known collectively as Nonlinear Dimensionality Reduction (NLDR), and we believe the utilization of NLDR approaches will allow a significant performance improvement for automated data analysis systems. In this report, we review the basic elements of NLDR techniques, we discuss the advantages of these techniques over more traditional approaches such as Principal-Component Analysis (PCA), and we outline an approach for utilizing NLDR to detect activities leading to the placement of IEDs based on airborne persistent surveillance video data.

15. SUBJECT TERMS
IED detection
Persistent surveillance

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified	UL	12	Frank Bucholtz / (202) 767-9342

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Executive Summary

Our ability to acquire data has, in many cases, far outstripped our ability to rapidly analyze that data. An important example is Persistent Surveillance (PS) which brings the capability to continuously monitor entire urban areas but which generates so much video data so rapidly that it is difficult for human evaluation to be performed reasonably, accurately, and quickly enough to be of actionable value. Of great utility in this case would be a computer-automated approach to identify and select only a small portion of the PS data for more careful scrutiny by an image analyst (IA). For example, an algorithm that rapidly identified a certain type of vehicular activity would be potentially useful for identifying activity leading to placement of roadside IEDs. That particular suspicious activity may be a more readily identifiable signature than the IED itself.

We believe that an important DoD goal should be the ***development of the tools and algorithms needed to implement an automated method for rapidly and accurately identifying suspicious activity and intent based on airborne persistent surveillance data***.

To achieve this goal we have identified a class of intelligent data reduction algorithms known collectively as *Nonlinear Dimensionality Reduction* (NLDR). These algorithms have had great success in a limited number of applications where traditional, linear techniques fail [1-14]. We believe there now exists an opportunity to focus further developments and refinements to the specific, important problem of identifying and classifying suspicious activity from video and image data.

Although recent progress in this area has been impressive, certain technical issues still remain that must be overcome before the technology can be utilized in a field-deployable system. Work to date has not considered the effects of noise, clutter, imperfect data, and truly high-complexity data sets. Therefore, we propose to examine two general issues: 1) the data quality issue, and 2) the data complexity issue. The quality issue involves the deleterious effects of nonideal data. For example, actual PS data may suffer from slow update rates, low spatial resolution, and occlusions due to buildings, trees, and vehicles. The effect of data deficiencies on the performance of NLDR algorithms must be thoroughly understood. The complexity issue refers to

Manuscript approved August 8, 2008.

the fact that, in principle, the effectiveness of NLDR approaches should be independent of data size and dimensionality. However, a quantitative investigation of NLDR performance on large data sets has never been performed.

Once these two issues are resolved, we believe that NLDR will provide a significant leap in the performance of automated data analysis systems.

Background

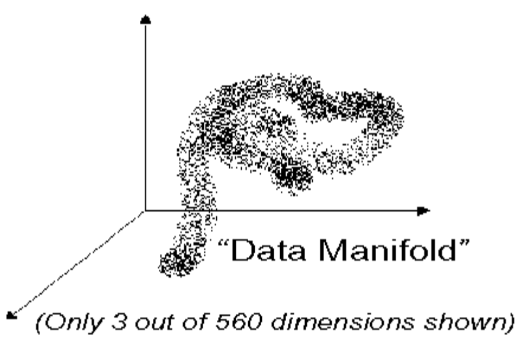
In this section we provide a brief introduction to image analysis approaches based on Nonlinear Dimensionality Reduction (NLDR) techniques. We refer the reader to several papers [1-14]. We emphasize that, although the focus in this discussion is Persistent Surveillance (PS) data from airborne platforms, the NLDR signal processing techniques will work on any data set including, for example, internet traffic, cellphone signals, or banking activity.

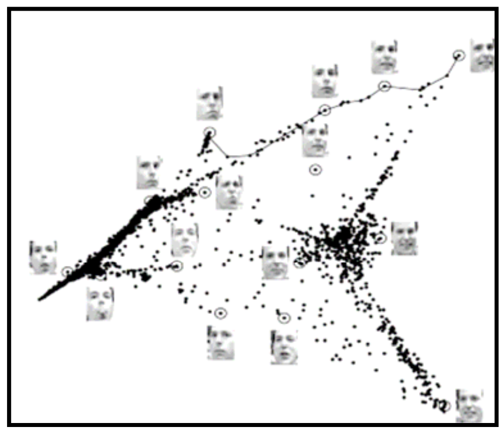
The purpose of any dimensionality reduction technique is to intelligently reduce the size of large, complex data sets so that information of interest can be identified and classified quickly and accurately while unnecessary information is ignored.   This is accomplished by transforming the original data to a smaller space that contains only information of interest. Extraneous information is discarded. This not only improves the process of information extraction but also significantly reduces computational effort. The exact process chosen to accomplish the reduction, however, can lead to vastly different results. We believe that by performing this reduction intelligently using NLDR, significant improvements in information extraction and classification are possible over conventional approaches.

We now introduce a key concept associated with any dimensionality reduction approach using an example from NLDR:

> Consider a large collection of photos of the type shown below of a face exhibiting different orientations and expressions [1]. Suppose we want to have a computer decide, when shown any one of the photos or even a new photo of the same person, whether the expression is a smile or a frown.

Each photo is an 8-bit gray-scale image comprising 20 x 28 pixels. Any one of the 560 pixels can have gray-scale values between 1 and 256. Let the gray-scale value of the j-th pixel $(1 \le j \le 560)$ in any one image be g_j. It is useful then to think of each image as a point in 560-dimensional space having position $(g_1, g_2, \cdots, g_{560})$ and to denote the value "560" as the "dimensionality" of the data (in this case, the collection of images). In general, each image will correspond to a distinct point in this 560-dimensional space. When all the points are plotted together, they typically form a well-defined geometric "object" referred to as a manifold (left-side picture in the figure below). As might be expected, similar facial expressions reside close together on the manifold. Hence, an approach for automatic





recognition of the type of facial expression is a purely geometric one: Designate regions on the manifold according to the general type of facial expression and, as each new image is encountered, see where the image resides on the manifold and classify the expression accordingly. However as the dimensionality of the data increases significantly, this approach can become costly in terms of both computation time and classification accuracy because ***large amounts of information that are not essential for classification and that represent noise or clutter are also processed***. An obvious question is "Do we really need all 560 dimensions to classify the facial expressions?" In most cases, the answer is NO. Provided it is done properly, research in NLDR has shown that the data dimensionality can be reduced significantly with only a small degradation in useful information. This effect is illustrated in

> the right-side plot in the figure above where the dimensionality of the space of images has been reduced from 560 to 2 using a NLDR algorithm called Locally Linear Embedding [1]. Hence, instead of working directly in 560-dimensional space, a computer-based "facial expression determination" algorithm would only need to analyze data in this 2-dimensional space! Both computation time and classification accuracy are significantly improved.
>
> It is important to note that the dimensionality reduction achieved here was not simply the elimination of 558 of the original data dimensions. **Instead, the data in all 560 original dimensions were used to compute an optimum 2-dimensional data space that retained most of the useful information in the original data.**

An important advantage of NLDR compared to conventional approaches concerns how the data are treated mathematically. Conventional approaches typically produce a new, smaller data space from <u>linear</u> combinations of the original data. One common example is the Principal Component Analysis (PCA) approach which seeks linear combinations of the original data axes along which the data shows highest variance, next-highest variance, etc. The assumption of linearity is a severe constraint since *there is no reason to believe that the key pieces of information to be extracted from the data are linearly separable from the noise and clutter*. NLDR approaches recognize this fact and allow the data to be <u>nonlinearly</u> related. The result is a data reduction approach that much more accurately captures the proper information relationships among the data thus allowing for accurate classification. A simple example is shown in Figure 1. In this simple example, the original data lives on a manifold known as the "Swiss roll" – a manifold shape that is particularly useful for illuminating the differences between linear and nonlinear approaches to dimensionality reduction. Here both PCA and LLE algorithms were applied to obtain a 2-D reduced dimensionality space. Both PCA and LLE mapped points that were close together in the original data to points that are close together in the reduced dimensionality space – a beneficial result. Unfortunately, PCA also maps points that are far away

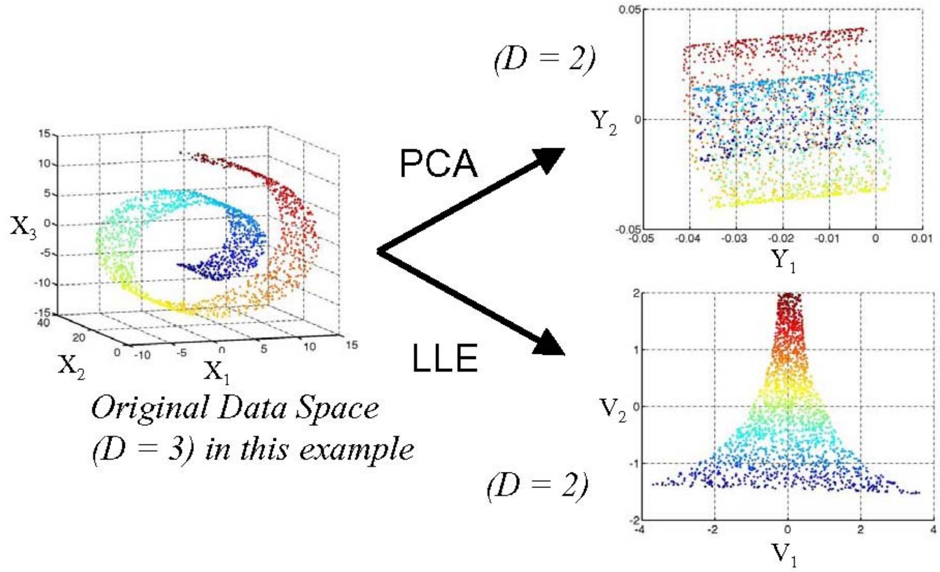


Fig. 1. *Comparison of conventional approach to data reduction (PCA) and one of the NLDR approaches, Local Linear Embedding (LLE). The PCA-based reduction cannot resolve the true relationship among the data points and the end result would be a large number of false alarms. LLE reduction preserves the correct information relationships among the data. Here, the PCA approach did NOT simply "squash" the original data onto the x1-x2 plane, rather it formed a linear combination of the x1-x2-x3 axes to obtain new y1-y2 axes depicted in the upper right plot.*

from each other in the original space (dark red and dark blue, for example) on top of each other in the reduced ($D = 2$) space. This will inevitably lead to confusion in the reduced space concerning the information relationship among these points. On the other hand, LLE clearly maintains the proper relationship among the red, yellow, and blue points in the reduced space and classification will be much more accurate in this case. This example serves as a useful analogy to the problem at hand, namely, the classification of suspicious behavior leading, for example, to IED placement. For example, the red points might encode the suspicious behavior of interest while the other colors may represent normal day-to-day activities. The ability to separate red points from other colors is analogous to separating suspicious activity from normal activity. Clearly, this is more rapidly accomplished in a reduced data space and more reliably accomplished in the reduced space resulting from the LLE compared to PCA.

It is important to note that the PCA approach did not simply "squash" the original 3-D

data down onto the plane depicted in Fig. 1 by the x_1 and x_2 axes. Instead, PCA found two distinct liner combinations of the original data space along which the data had the largest and next-largest variance. By comparison, LLE seeks to preserve the proper geometric relationships among *local* groups of points on the manifold. By focusing on local geometric properties only, the LLE approach can handle data that is globally nonlinear. The same is true of other NLDR approaches, as well.

NRL Approach to Automated Analysis of Persistent Surveillance Data

In this section we outline our overall approach for automated processing of persistent surveillance (PS) data with an emphasis on identifying suspicious behavior corresponding to the placement of roadside IEDs.

The NRL approach is shown schematically in the Fig. 2. This example is based on actual PS video data obtained during an NRL field test in June 2007 together with simulated movements of ground vehicles. We begin with a PS video file containing multiple, sequential images of a fixed geographical region taken over some time period. Here we used 100 frames each having 577 x 952 pixels. The dimensionality of this original data space is, therefore, D = 54,930,400. The steps in the automated PS analysis would be:

1) Using existing image registration techniques, the raw image data are registered to obtain a video mosaic. (In this test example, we used one fixed frame but several image registration techniques are available [15-17])

2) Existing tracking algorithms are then employed to extract the tracks of all moving objects. In this test example, we simulated vehicle tracks along actual roads. Each track, comprising the (x,y,z) location of the vehicle in 100 frames, now represents a point in D = 300 dimensional space. Here the red tracks correspond to the behavior of interest – namely, vehicles that stopped momentarily along the roadside. The blue colored tracks represent vehicles moving without stopping.

3) An NLDR algorithm (LLE in this case) is applied to reduce the D = 300 dimensional track data to a 3-D data space for classification purposes. The two types of activities are clearly separated on the manifold in the 3-D reduced space. Hence, the classification is now performed in a 3-D data space compared to the original D = 54,930,400 data space!

4) Each track would then be analyzed to determine where it occurred on the manifold in reduced dimensionality space. If it occurred on a "benign" portion of the manifold it would be ignored. If it occurred on a "threat" portion of the manifold, an IA could be alerted to examine that particular track more closely or the event could be flagged for forensic purposes.

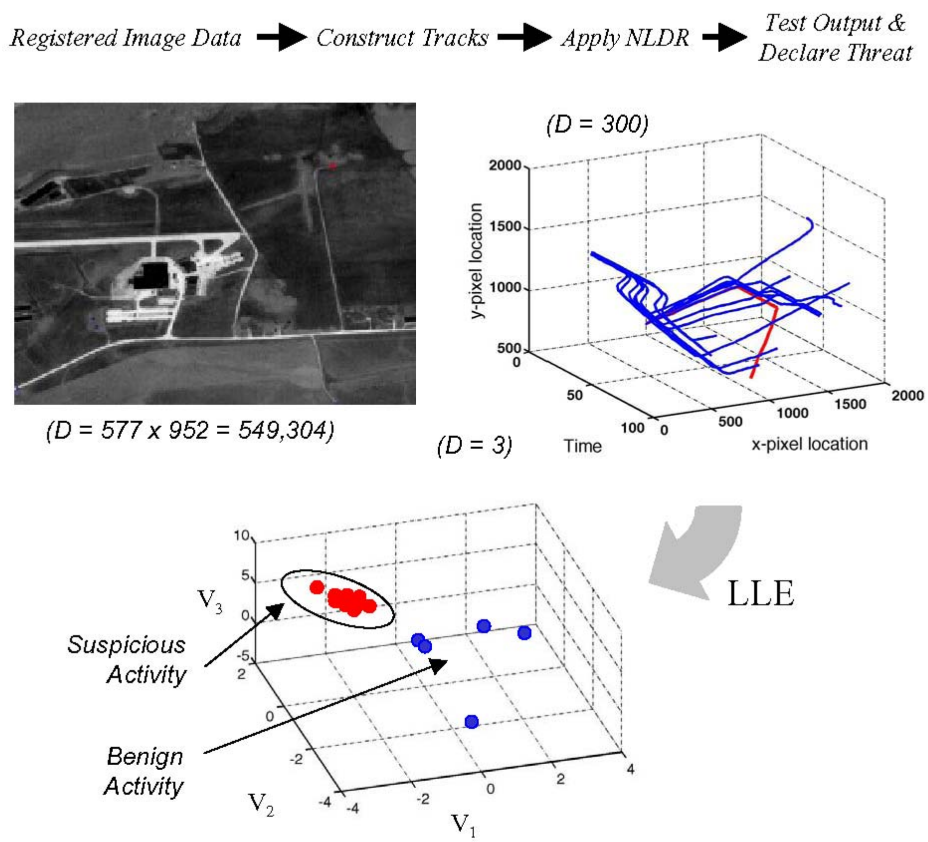


Fig. 2. *NRL approach to data dimensionality reduction for persistent surveillance data. After registration of sequential images, vehicle tracks are extracted and processed using an NDLR algorithm such as LLE. Once a manifold is established in reduced-dimensionality space, the location of each new track determines whether the activity represented by the track is threatening or benign.*

Hence, the NRL approach employing NLDR can be summarized as follows. We begin with a very large, very complex data set comprising a video file obtained from a persistent surveillance platform. Instead of analyzing the raw data directly, we first extract vehicle tracks, and then apply an NLDR technique to display all the data of interest in just three dimensions. **That is, the large, complex raw data set has been intelligently reduced to a set of points in just three dimensions that can be analyzed quickly and robustly by a computer.**

We compared the performance of LLE against PCA for the example. The results are shown below in Fig. 3. Clearly the LLE approach yields superior performance since all tracks corresponding to stopped vehicles are clearly separated from those that did not. Conventional PCA confused tracks that pause from those that do not.

It should be stressed at this point that although we have chosen to focus on tracks that pause by the roadside, ***the algorithm can search for any type of activity of interest.*** Different activities will lie on different parts of the reduced data manifold. For example, one might choose to search for all vehicle tracks that pass a certain area at a certain velocity. Additionally, as was mentioned earlier, it does not matter what type of data are used in the analysis. The data could come only from vehicle tracks, as was done here, but it could include other pieces of information such as time of day, data from street-level cameras, known addresses and locations, etc.

Another advantage of the proposed approach is that the experience and intuition of the IAs can be incorporated into the algorithms in the form of a threat library. This library can (and likely will) change as the nature of the threat changes, thus rendering the approach highly flexible, adaptive to new activities of interest and yielding significantly fewer false alarms than more conventional approaches.

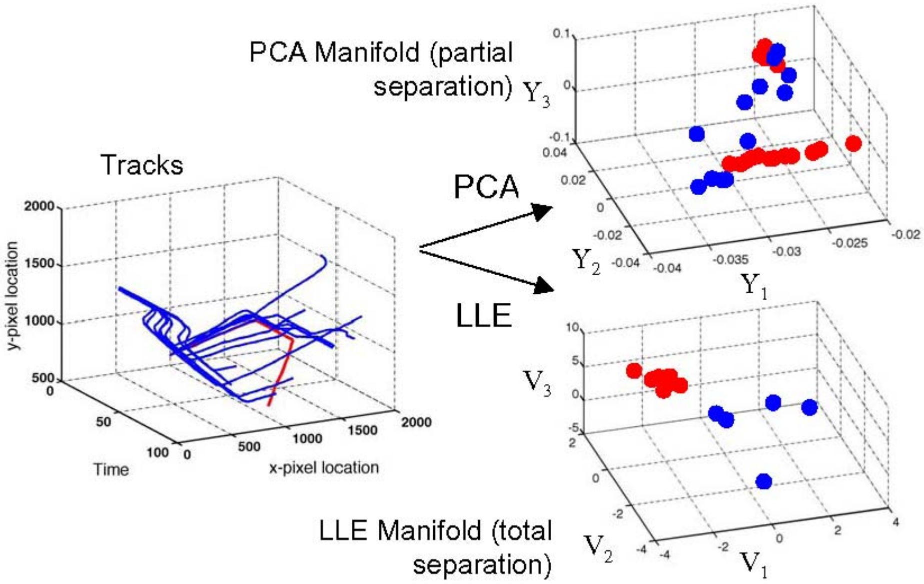


Fig.3.*Comparison of the linear PCA approach with the nonlinear LLE approach for the simulated track data in the NRL experiment. The manifold resulting from LLE clearly separates threatening (red) tracks from benign (blue) tracks allowing for proper classification by an automated system. The PCA approach mixes the two types of activities making separation difficult.*

Summary

Algorithms based on nonlinear dimensionality reduction (NLDR) techniques show great promise for enabling the rapid analysis of large, complex data sets. An important application for such analysis techniques will be for persistence surveillance video data and, in particular, for the detection of suspicious activities such as activity leading to the placement of IED devices.

Acknowledgements

The authors acknowledge helpful discussions with N.J. Frigo and J.V. Michalowicz.

References

[1] S.T. Roweis and L.K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” Science **290** 2323-2326 (2000).

[2] R. R. Coifman, S. Lafon, A.B. Lee, M. Maggioni, B. Nadler, F. Warner, and S.W. Zucker, “Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps,” Proc. Nat. Acad. Sci. **102** (21) 7426-7431 (2005).

[3] B. Nadler, S. Lafon, R. R. Coifman, and I.G. Kevrekidis, “Diffusion maps, spectral clustering and reaction coordinates of dynamical systems,” arXiv:math/0503445v1 (2005).

[4] J.B. Tenenbaum, V. de Silva, and J.C. Langford, ”A global geometric framework for nonlinear dimensionality reduction,” Science, **290** 2319-2323 (2000).

[5] C.M. Bachmann, T.L. Ainsworth, and R.A. Fusina, “Exploiting manifold geometry in hyperspectral imagery,” IEEE Trans. Geosci. Remote Sens., **43** (3) 441-454 (2005).

[6] C.M. Bachmann, T.L. Ainsworth, and R.A. Fusina, “Improved manifold coordinate representations of large-scale hyperspectral scenes,” IEEE Trans. Geosci. Remote Sens., **44** (10) 2786-2803 (2006).

[7] D.L. Donoho and C. Grimes, “Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data,” Proc. Natnl. Acad. Sci. **100** (10) 5591-5596 (2003).

[8] K. Fukunaga and D.R. Olsen, “An algorithm for finding intrinsic dimensionality of data,” IEEE Trans. Comp. **c-20** (2) 176-183 (1971).

[9] M. Belkin and P. Niyogi, “Towards a theoretical foundation for Laplacian-based manifold methods,” Proc. Learning Theory, **3559** 486-500 (2006).

[10] M. Belkin and P. Niyogi, “Laplacian eigenmaps for dimensionality reduction and data representation,” Neural Comp. **15** (6) 1373-1396 (2003).

[11] Z. Zhang and H. Zha, “Principal manifolds and nonlinear dimension reduction via Local Tangent Space Alignment, “ *Technical Report*, CSE-02-019, Department of Computer Science & Engineering, Pennsylvania State University.

[12] Z. Zhang and H. Zha, “Principal manifolds and nonlinear dimension reduction via tangent space alignment, “ SIAM J. Sci. Comp. **26** (1) 313-338 (2004).

[13] Z. Zhang and H. Zha, “Continuum isomaps for manifold learnings,” Comput. Stat. Data Anal., **52** (1) 184-200 (2007).

[14] P.W. Jones, M. Maggioni and R. Schul, “Manifold parameterizations by eigenfunctions of the Laplacian and heat kernels,” Proc. Natnl. Acad. Sci., **105** (6) 1803-1808 (2008).

[15] L.G. Brown, ”A survey of image registration techniques,” ACM Computing Surveys,” **24** 326-376 (1992).

[16] B. Zitova and J. Flusser, “Image registration methods: A Survey,” Image Vision Computing, **21** 977-1000 (2003)

[17] S. Blackman and R. Popoli, Design and Analysis of Modern Tracking Systems, Artech House (1999)